DOCUMENT RESUME

ED 228 415 CE 035 509

AUTHOR Trafton, Richard S.; Garrison, William T.
TITLE Male-Female Differences in Cognitive Abilities in Various Age Groups. Technical Report 1982-3.
INSTITUTION Johnson O'Connor Research Foundation, Boston, MA.
PUB DATE Jul 82
NOTE 13p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Age Differences; *Cognitive Ability; Cognitive Measurement; Comparative Analysis; Educational Research; Females; Males; *Sex Differences

ABSTRACT

The research question addressed was whether sex differences in cognitive ability are smaller in younger age groups than in older age groups. A battery of 12 aptitude tests measuring visualization, spatial orientation, perceptual speed, associative memory, number memory, eidetic memory, spatial scanning, visual memory, and flexibility of closure were administered to a sample of 1,011 females and 1,226 males ranging in age from 14 to 63 years of age. Mean male and female performances on each test were compared for 22 age groups. Examination of mean differences for age did not support a suggested trend toward smaller differences between males and females on measures of cognitive ability. The results were consistent with previous research on sex differences. Males outperformed females on spatial tests; females outperformed males on the perceptual speed measure and the test of associative memory. (Three data tables are included.) (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED228415

MALE-FEMALE DIFFERENCES IN COGNITIVE ABILITIES IN VARIOUS AGE GROUPS

by

Richard S. Trafton
William T. Garrison

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

JOHNSON O'CONNOR RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

K Bowker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Technical Report 1982-3

July, 1982

CE035505

Male-Female Differences in Cognitive Abilities in Various Age Groups

Richard S. Trafton and William T. Garrison

SUMMARY: A battery of twelve aptitude tests measuring Visualization, Spatial Orientation, Perceptual Speed, Associative Memory, Number Memory, Eidetic Memory, Spatial Scanning, Visual Memory, and Flexibility of Closure was administered to a sample of 1011 females and 1226 males ranging in age from 14 to 63 years. Mean male and female performance on each test were compared within 22 age groups. The research question addressed was whether sex differences in cognitive ability are smaller in younger age groups than in older age groups. Examination of mean differences by age revealed no support for such a trend. Observed sex differences were consistent with those reported in the literature, except that no sex difference was found on Flexibility of Closure.

INTRODUCTION

In the past 20 years, much has been written concerning the differences between males and females on various cognitive abilities. It has generally been suggested that males, on the average, tend to outperform females on measures of spatial ability; females, on the average, tend to outperform males on measures of verbal ability (Maccoby & Jacklin, 1974; Petersen & Wittig, 1979). Concerning spatial tasks, it has been suggested that males outperform females on measures of what has traditionally been called Visualization (Vz), while there is essentially no difference between males and females on measures of Spatial Orientation (S) (Clements, Note 1). The interpretation is that the greater complexity of the Vz tasks, and the sequential processing they require, produce the sex difference. That is, males and females are equally capable on the two-dimensional, simple rotational or change of perspective tasks, but males are better at three-dimensional, multi-step procedures (Wattanawaha, reported in Clements, Note 1). Maccoby and Jacklin (1974) make this distinction in terms of analytic versus nonanalytic visualization, where analytic visualization involves some form of verbal mediation, while nonanalytic visualization does not. Petersen and Wittig (1979) suggest that timing of tests (as opposed to simply looking at number of correct responses) increases the likelihood of tapping nonanalytic visualization. The sex differences, favoring females, which appear on certain verbal tasks are less easily interpreted. However, this difference does appear on a variety of measures, including some related to verbal memory (Maccoby and Jacklin, 1974).

It has recently been suggested that there is a trend toward fewer or smaller sex differences on measures of spatial ability. Jacklin (1979) reports finding small or no sex differences on spatial measures using younger (more recent) samples. Although she does not actually suggest similar reductions in male-female differences on verbal measures, she does allude to a trend toward smaller differences in cognitive abilities in general. Presumably, these would include verbal abilities.

The purpose of the present study is to determine the extent to which there is a trend toward smaller differences between males and females on

measures of cognitive ability. This will be accomplished by examining sex differences at various ages, across a wide age range. If there is in fact a trend toward smaller sex differences, younger groups of males and females should show smaller differences on average than older groups. Furthermore, by studying a variety of cognitive abilities it is possible that some inferences may be drawn as to the nature of differences that are observed.

METHOD

Subjects. Participants in the study were clients of the Johnson O'Connor Research Foundation, an aptitude testing and research organization. All clients paid a fee to take a battery of tests to help them in vocational planning. A total of 1,011 females and 1,226 males (45% and 55% of the total sample, respectively) took part in the testing. Age of the clients ranged from 14 to 63 (mean = 23 years, standard deviation = 8.8 years). All geographic regions of the country, with the exception of the Pacific northwest, were represented in the sample. The average female in the sample (24.1 years) was older than the average male (22.2 years) ($t = 5.23$, $p < .001$). It was felt that this difference would not adversely affect any analyses.

Instruments. The following twelve tests were used in the present study. Asterisks indicate tests published by the Johnson O'Connor Research Foundation. The remainder of the tests are drawn from the Kit of Factor Referenced Cognitive Tests (Ekstrom, French, Harman & Dermen, 1976) and are explained in detail in the Kit manual. Because of time constraints imposed by the large number of tests administered, only the first of the two equivalent halves of the tests from the Kit were used.

Wiggly Block* (Worksample 4,5,6 AC). This is a block-assembly task with four individually timed trials with increasing levels of difficulty (6, 9 and 12 pieces). The test measures Vz. Reliability = .73.

Paper Folding* (Wks. 622 FA). This is a derivation of the Thurstone Punched Holes Test, but without the multiple-choice response format. Responses to the ten items are individually timed and are made by placing chips on a grid to indicate where the holes will be when the paper is unfolded. The test measures Vz. Reliability = .83.

Form Board. This is a marker test for Vz. Reliability = .68.

Surface Development. This test is also a marker for Vz. Reliability = .83.

Graphoria* (Wks. 613 CI). This is a number-checking task similar to the Number Comparison Test in the Kit. The test measures Perceptual Speed (P). Reliability = .90.

Object Observation* (Wks. 206). This is a test of eidetic memory. The subject is shown a photograph of twenty common objects and allowed to study it for one minute. Subsequent photographs are shown, interspersed with brief study periods of the original, and the subject is asked to indicate what changes have occurred (objects moved, removed, etc.). Reliability = .62.

Number Memory* (Wks. 165). This is not a memory span test, but measures memory skills related to grouping (chunking). The subject is presented, visually, with a sequence of eight six-digit numbers, with a brief study period between numbers. He or she is then asked to recall as many of the numbers as possible. There are four trials using the same set of numbers. Reliability = .91.

Silograms* (Wks. 376). This is a test of associative memory. A list of English words is paired with a list of (pronounceable) nonsense words. The pairs are presented visually. Then, the nonsense words are presented with the subject being required to provide the appropriate English associate. Reliability = .92.

Shape Memory. This is a marker test for the Visual Memory (MV) factor from the Kit. Reliability = .52.

Cube Comparison. This is a marker test for S. Reliability = .63.

Map Planning. This is a marker test for the Spatial Scanning (SS) factor from the Kit. Reliability = .63.

Hidden Figures. This is a marker test for the Flexibility of Closure (CF) factor from the Kit. This test has also been used as an indicator of Field Independence/Dependence. Reliability = .68.

RESULTS

The male and female means for each test were compared, collapsing across age groups, using *t* tests. If there were no trend toward smaller sex differences, these comparisons should serve as the best estimate of the actual population differences. The results of these *t* tests are presented in Table 1. In addition to the values of *t* and *p*, Table 1 shows the "difference ratios," defined as the female mean minus the male mean, divided by the pooled standard deviation (the square root of the weighted average of the within-sex variances).

The differences are statistically significant ($p < .01$) for all tests except Number Memory and Hidden Figures. Because of the large samples used, even very small differences could be found significant. However, most of the statistically significant differences are relatively large in terms of standard deviation units. For example, the Vz measures have differences averaging approximately .30 units (Wiggly Block = .33, Paper Folding = .27, Form Board = .33, Surface Development = .30). Assuming normal distributions, about 38% of females would score above the male median under these conditions. Similarly, the difference on Graphoria (P) is .44 units, favoring females. In this instance, 33% of males would score above the female median. Similar statements can be made regarding the other tests by treating the differences expressed in standard deviation units as normal deviate scores.

The comparisons in Table 1 do not, however, address the question of age trends in the differences. To this end, subjects were grouped by age

Table 1
Sex Differences for All Ages Combined

Test	Mean Female	Mean Male	Higher Sex	Difference Ratio[a]	*t*	*p*
Wiggly Block[b]	3.26	2.43	M	.33	7.83	.000
Paper Folding	16.33	19.89	M	-.27	6.30	.000
Form Board	8.16	9.52	M	-.33	7.69	.000
Cube Comparison	10.64	11.62	M	-.25	5.74	.000
Map Planning	9.96	10.46	M	-.13	3.08	.002
Surface Development	15.08	17.47	M	-.30	6.93	.000
Silograms	40.98	32.37	F	.52	12.59	.000
Graphoria	163.19	152.57	F	.44	10.56	.000
Object Observation	67.15	63.62	F	.28	6.82	.000
Number Memory	78.37	75.85	F	.09	2.23	.026
Shape Memory	12.01	11.47	F	.22	5.00	.000
Hidden Figures	5.94	5.67	F	.07	1.56	.118

[a] The difference ratio is the female mean minus the male mean, divided by the pooled standard deviation (square root of the weighted mean of the within-sex variances).

[b] This test is scored on time, so a lower raw score indicates better performance.

such that no age group had fewer than ten males or ten females. In relatively few groups was the number of males or females less than 20. The number of subjects in each group is given in Table 2. Twenty-two age groups ranging from 14-15 years to 50 years and over were produced. The difference ratio was calculated separately for each age group, in the same way as for Table 1. The difference ratios are affected by sample size in that smaller samples will produce less stable estimates of the mean (i.e., the standard error of the mean is inversely related to the square root of the sample size), and therefore less stable estimates of the difference ratio. However, the estimates are unbiased. Difference ratios for all age groups are shown in Table 2.

If there is a temporal trend toward smaller sex differences, it should be manifested in a correlation between the difference ratios and age. The correlation of age with difference ratio for each test is presented in Table 2. Two tests produced statistically significant correlations between age and difference ratio: Surface Development and Hidden Figures.

On tests where females tend to score higher (as indicated by positive difference ratios), a statistically significant positive correlation would support an inference of a trend over time toward smaller sex differences. For tests on which males tend to score higher (negative difference ratios), an inference of such a trend would be supported by a statistically significant negative correlation. The negative correlation for Surface Development indicates that the male advantage for older groups on this test is larger than the male advantage for younger groups. The result for Hidden Figures suggests that the trend is from a male advantage in older groups to a female advantage in younger groups.

Table 2

Difference Ratios by Age Group

Age	Test												N^a	
	WB[b]	PF	SLGM	Graph	ObObs	# Mem	SM	FB	CC	MP	SD	HF	Female	Male
14-15	-1.63	-.22	.45	.31	.29	.19	-.15	-.46	-.26	.00	-.07	-.16	43	45
16	-.26	-.30	.25	.28	.29	.19	.11	-.20	.00	.00	.04	.04	70	77
17	-.31	-.34	.60	.49	.24	.12	.41	-.15	-.20	-.09	-.22	.36	120	180
18	-.33	-.31	.70	.43	.35	.12	.38	-.31	-.07	-.06	-.24	.14	91	128
19	-.17	-.36	.83	.51	.06	.27	.43	-.21	-.04	-.20	-.01	.05	83	100
20	-.44	-.11	.20	.36	.05	-.16	.05	-.43	-.31	-.16	-.43	-.04	57	110
21	-.28	-.06	.63	.44	.49	-.08	.24	-.62	-.57	-.02	-.41	.14	30	60
22	-.05	.06	.86	.63	.22	.05	.62	-.14	-.01	-.04	-.14	.50	38	36
23	-.39	-.32	.76	.60	.23	.04	.47	-.62	-.64	-.23	-.39	-.40	35	26
24	-.28	-.03	.43	-.18	.54	-.11	.24	-.24	-.33	-.18	-.17	.13	35	29
25	-.10	-.02	.43	.89	.58	.08	-.07	.07	-.19	.14	-.08	.50	19	26
26	-.24	-.20	.25	.52	.32	-.17	.13	.37	-.34	-.35	-.38	.33	15	31
27	-.37	.06	.88	.92	.10	.27	.15	-.18	-.18	-.03	-.48	.39	11	17
28	-.36	.08	.90	.13	.54	-.03	.24	.10	-.06	-.23	-.02	-.13	19	15
29	-.23	-.25	.84	.11	.77	.18	.00	-.14	-.02	.03	-.07	-.07	15	20
30-31	-.60	-.73	.77	.20	.31	.05	-.24	-.58	-.45	-.51	-.56	-.34	31	20
32-33	-.12	-.11	.23	.04	.55	.12	.17	-.29	-.15	-.19	-.11	.08	34	24
34-35	-.49	-.45	.37	.60	.63	.36	.34	-.92	-.19	-.42	-.71	-.09	33	24
36-39	-.53	-.03	.64	.12	.75	-.35	.15	.17	-.09	-.02	-.17	-.21	44	21
40-43	-.64	-.34	.31	.50	-.05	.19	.08	-.66	-.51	-.46	-.79	-.13	27	18
44-49	-.38	-.50	.37	-.26	.22	-.10	.11	-.60	-.23	-.74	-1.18	-.56	36	22
50+	-.27	-.56	-.04	.96	.30	.00	.02	-.54	-.43	-.48	-.72	-.22	15	20
Corr. of DR with Age	.09	-.35	-.36	-.06	.11	-.20	-.26	-.25	-.22	-.18	-.70**	-.49*		

* $p < .05$ ** $p < .01$

[a] Sample sizes vary among tests.

[b] Signs reversed for consistency with other tests. A negative difference ratio indicates that males score higher.

It is important to note that even though the overall sample size is large, the power associated with the tests of statistical significance of the correlations is still a function of the number of bivariate data points, that is, age/difference-ratio pairs. With 22 "observations" (determined by the number of age groups), the power of the test for a modest correlation ($\rho = .20$) is only .15; for a strong correlation ($\rho = .50$), it is .69. So the probability of not finding a statistically significant result even if the correlation is moderate or large is higher than many researchers would desire.

Another perspective on these data is provided by examining *t*-test comparisons of male and female means for each test in each age group. These comparisons are presented in Table 3. Because of the large number of statistical tests performed, the experiment-wide alpha level is quite high ($> .99$). In fact, thirteen of the tests would be expected to be significant (at $\alpha = .05$) by chance alone. Of the 264 comparisons, 65 are statistically significant, clearly more than chance expectations. However, it is impossible to know which significant results represent the errors. It is most likely, though, that the Type I errors would be found in anomalous cases, rather than elements in sets.

With these considerations in mind, examination of Table 3 shows a relatively even distribution of statistically significant differences across age groups, with the exception of Silograms and Shape Memory, which show a greater number of significant differences for younger groups, and Surface Development, which shows more significant differences for older groups. This result is consistent with the outcome of the correlational analysis.

DISCUSSION

The trend toward smaller sex differences, suggested by Jacklin (1979), if present, should have been manifested by correlations between age and difference ratios, or by different patterns of between-sex *t*-tests within age groups, for the various age groups studied. Neither of these outcomes was generally evident in the present study. The results were in fact consistent with previous research on sex differences (cf., Maccoby & Jacklin, 1974; Wittig & Petersen, 1979). Males outperformed females on spatial tests; females outperformed males on the perceptual speed measure and the test of associative memory.

Contrary to suggestions presented by Clements (Note 1), males showed about one-fourth of a standard deviation advantage on the Cube Comparison Test, a measure of spatial orientation (S). Although this advantage is not quite as large as that found for the visualization measures, it is still substantial. No support is found in the present study for the suggestion that the male advantage on spatial ability is found only on three-dimensional, sequential tasks. Also without support is the inference by Jacklin (1979) that sex differences in general are decreasing. The general trend appears to be for females to do better relative to males at younger ages than at older ages. Thus, on tests where females do better than males, the sex difference appears to be increasing over time.

Table 3

Results of t-tests by Age Groups

Age	Test											
	WB	PF	SLGM	Graph	ObObs	#. Mem	SM	FB	CC	MP	SD	HF
14-15	-3.26**	-1.06	2.13*	1.50	1.38	.90	- .72	-2.15*	- .61	- .02	- .36	- .76
16	-1.70	-1.90	1.61	1.84	1.90	1.25	.68	-1.24	- .02	0	.23	.21
17	-2.69**	-2.98**	5.47**	4.44**	2.20*	1.03	3.47**	-1.25	-1.68	- .74	-1.88	2.95**
18	-2.52*	-2.41*	5.49**	3.35**	2.80**	.91	2.80**	-2.37*	- .50	- .40	-1.81	1.02
19	-1.17	-2.53*	5.63**	3.46**	.43	1.88	2.90**	-1.43	- .27	-1.38	- .04	.34
20	-2.48*	- .63	1.22	2.19*	.32	-1.02	.34	-2.68**	-1.93	-1.03	-2.50*	- .24
21	-1.19	- .24	3.10**	2.10*	2.35*	- .38	1.02	-2.72**	-2.62*	- .07	-1.77	.63
22	- .21	.23	3.77**	2.76**	.97	.21	2.66**	- .61	- .05	- .18	- .61	2.17*
23	-1.51	-1.18	2.95**	2.29*	.90	.56	1.80	-2.39*	-2.47*	- .89	-1.51	-1.52
24	-1.16	- .12	1.74	- .72	2.17*	- .44	.91	- .94	-1.30	- .72	- .67	.53
25	- .32	- .05	1.43	2.95**	1.95	.28	- .23	.24	- .62	.45	- .25	1.67
26	- .78	- .63	.80	1.69	1.02	- .56	.42	1.18	-1.10	-1.12	-1.22	1.07
27	- .96	.15	2.29*	2.40*	.25	.68	.39	- .45	- .48	- .07	-1.22	.99
28	-1.05	.25	2.61*	.39	1.57	- .08	.70	.30	- .16	- .68	- .07	- .36
29	- .68	- .74	2.48*	.33	2.28*	.54	.01	- .42	- .05	.08	- .21	- .19
30-31	-2.09*	-2.50*	2.69**	.71	.49	.17	- .84	-2.02*	-1.56	-1.75	-1.92	-1.17
32-33	- .47	- .40	.87	.17	2.06*	.46	.63	-1.04	- .57	- .71	- .42	.27
34-35	-1.83	-1.69	1.39	2.26*	2.38*	1.36	1.27	-3.39**	-1.15	-1.56	-2.62*	- .33
36-39	-2.16*	- .06	1.04	.26	2.87**	-1.42	.45	.66	- .37	.78	- .68	- .86
40-43	-2.27*	-1.13	1.43	1.74	- .17	.67	.29	-2.30*	-1.79	1.74	-3.00**	- .43
44-49	-1.44	-1.88	1.43	- .99	1.03	- .41	.41	-2.26*	- .90	-2.79**	-4.57**	-2.16*
50+	- .86	-1.76	- .11	3.04**	.94	- .01	.07	-1.66	-1.32	1.51	-2.17*	- .65

* $p < .05$ ** $p < .01$

Note: Sign of *t* indicates which sex scores higher: negative = males higher, positive = females higher.

One of the tests, Surface Development, did offer results supporting a temporal trend toward a decrease in the size of the sex difference. Although the correlation of age with difference ratio (.70) is large, the other three tests of the Vz factor did not show a statistically significant relationship. Therefore, conclusions based on this test alone should be attenuated. The results for Hidden Figures are more interesting. This was the test with the smallest sex difference overall. The trend is from a male advantage in the older groups to a female advantage in the younger groups. If this test is a function of the general dimension of Field Independence/Dependence, there may be some social phenomenon producing this trend. However, interpretations of this test should be made with care owing to the suggestion of underlying construct differences for males and females (Garrison and Trafton, Note 2).

The present study has the advantage of using tests with known factorial content (with the possible exception of Hidden Figures). It may be that the purported trends observed by other researchers in this area are the result of using tests which are less factorially distinct. For example, much of the work of Fennema and Sherman (1977) is based on the Space Relations subtest of the Differential Aptitude Tests. This test is not a pure measure of any commonly identified spatial ability.

The possibility always remains that the results reported here are sample-specific. Even though the overall sample was quite large (giving a very stable estimate of the average sex difference if no age trends are assumed), the sample sizes for each age group were not. This issue will be resolved by replication. Data are now being gathered to allow such replication.

REFERENCE NOTES

1. Clements, M.A. Spatial ability, visual imagery, and mathematical learning. Paper presented at the meeting of the American Educational Research Association, Los Angeles, April, 1981.

2. Garrison, W.T. and Trafton, R.S. Components of performance on the Hidden Figures Test (Technical Report 1981-3). Boston: Johnson O'Connor Research Foundation, 1981. Paper presented at the meeting of the American Educational Research Association, Los Angeles, April, 1981.

REFERENCES

Ekstrom, R.B., French, J.W., Harmon, H.H., and Derman, D. Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, N.J.: Educational Testing Service, 1976.

Fennema, E. and Sherman, J. Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Education Research Journal, 1977, 14, 51-71.

Jacklin, C.N. Epilogue. In M.A. Wittig and A.C. Petersen (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

Maccoby, E., and Jacklin, C.N. The psychology of sex differences. Stanford: Stanford University Press, 1974.

Peterson, A.C. and Wittig, M.A. Sex-related differences in cognitive functioning: An overview. In M.A. Wittig and A.C. Petersen (Eds.), Sex-related differences in cognitive functioning. New York: Academic Press, 1979.

Wittig, M.A. and Petersen, A.C. Sex-related differences in cognitive functioning. New York: Academic Press, 1979.